AD-A104 372 CHICAGO UNIV IL CENTER FOR DECISION RESEARCH F/G 5/10
PREDICTION, DIAGNOSIS, AND CASUAL THINKING IN FORECASTING.(U)
SEP 81 H J EINHORN, R M HOGARTH N00014-81-K-0314

UNCLASSIFIED NL

1 of 1
AD
A104372

END
DATE
FILMED
10-81
DTIC

[illegible] Vita
[illegible] exco-
[illegible] latur

Prediction, Diagnosis, and Causal
Thinking in Forecasting*

Hillel J. Einhorn   Robin M. Hogarth
Center for Decision Research
Graduate School of Business
University of Chicago
1101 East 58th Street
Chicago, Illinois 60637
(312) 753-4957

*Support for this work was provided by a contract from the Engineering Psychology Program, Office of Naval Research.

Journal of Forecasting, in press.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 1	2. GOVT ACCESSION NO. AD-A104372	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) PREDICTION, DIAGNOSIS, AND CAUSAL THINKING IN FORECASTING		5. TYPE OF REPORT & PERIOD COVERED Interim Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Hillel J. Einhorn [redacted] Robin M. Hogarth		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0314
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Decision Research, Univ. of Chicago 1101 East 58th Street Chicago, Illinois 60637		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 197-071
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217		12. REPORT DATE Sep[redacted]81
		13. NUMBER OF PAGES 33
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

To be published in, The Journal of Forecasting, 1982.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

causal judgment
cues to causality
diagnostic processes
forecasting
forward and backward inference
model formation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

While forecasting involves forward/predictive thinking, it depends crucially on prior diagnosis for suggesting a model of the phenomenon, for defining "relevant" variables, and for evaluating forecast accuracy via the model. The nature of diagnostic thinking is examined with respect to these activities. We first consider the difficulties of evaluating forecast accuracy without a causal model of what generates

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

413544

outcomes. We then discuss the development of models by considering how attention is directed to variables via analogy and metaphor as well as by what is unusual or abnormal. The causal relevance of variables is then assessed by reference to probabilistic signs called "cues to causality." These are: temporal order, constant conjunction, contiguity in time and space, number of alternative explanations, similarity, predictive validity, and robustness. The probabilistic nature of the cues is emphasized by discussing the concept of spurious correlation and how causation does not necessarily imply correlation. Implications for improving forecasting are considered with respect to the above issues.

Imagine that you lived several thousand years ago and belonged to a tribe of methodologically sophisticated cave-dwellers. Your methodological sophistication is such that you have available to you all present day means of the methodological arsenal--details of the principles of deductive logic, probability theory, access to computational equipment, etc. However, your level of substantive knowledge lags several thousand years behind your methodological sophistication. In particular, you have little knowledge about physics, chemistry or biology. In recent years, your tribe has noted an alarming decrease in its birth-rate. Furthermore, the tribe's statistician estimates that unless the trend is shortly reversed, extinction is a real possibility. The tribe's chief has accordingly launched an urgent project to determine the cause of birth. You are a member of the project team and have been assured that all means, including various forms of experimentation with human subjects, will be permitted to resolve this crucial problem.

The above story illustrates the following points:

(1) The ultimate goal of forecasting is to provide guidance for taking action. Therefore, forecasting is intimately tied to decision making and should be evaluated within this context. For example, although the tribe's statistician may use a time series model of past birth rates to predict the downward trend, this is of limited usefulness to the tribe's leaders. Indeed, one might find small consolation in the fact that the date of the tribe's extinction can be precisely predicted. Clearly, the need to take appropriate action to change the process is of paramount concern.

(2) While much attention in the forecasting literature is devoted to forecast accuracy (see Hogarth & Makridakis, 1981, for a review), this focus may be misplaced. In particular, when viewed within the broader context of how forecasts affect actions, and how both affect outcomes, the criterion of

forecast accuracy is problematic. In order to illustrate this, consider the following three situations and the relevance of forecast accuracy in each: (a) A weather forecaster predicts that a hurricane will come ashore at a certain time and place and actions contingent on the forecast are taken (e.g., boarding up property, evacuation). Since these actions do not affect the hurricane itself, the relevance of forecast accuracy as a criterion is obvious and its evaluation relatively simple. However, even in those cases where one cannot typically exert control over the process (as when large physical systems are involved), actions can sometimes be taken to affect the process; consider, for example, "seeding" hurricanes (Howard, Matheson, & North, 1972); (b) Economic forecasters predict a recession next year and the government reacts by taking action to stimulate the economy. Clearly, action is taken to intervene in the process and the resulting outcomes are a joint function of the factors on which the prediction is based and the action taken. Under these circumstances, the meaning of forecast accuracy is ambiguous; in fact, one can imagine that forecast accuracy is <u>not</u> desirable. For example, the more effective the action is in changing the process, the less accurate the initial forecast will be; (c) People in a small town hear a rumor that the banks are about to fail. They think that if this forecast is accurate, they had better withdraw their money as fast as possible. Accordingly, they go to the banks to close their accounts (those skeptical of the forecast see many people withdrawing money and either take this as a sign that the rumor is true or foresee the consequences of waiting too long, thus joining the crowd in either case). By the end of the day the banks have failed, thereby confirming the rumor. This case differs from (b) in that the actions taken lead to confirmation of the predictions (demonstrating the so-called "self-fulfilling" prophecy effect). The perniciousness of these cases is that awareness of the effect of actions on outcomes is low and can lead to overconfidence in forecasts that are of low or even zero accuracy (see Einhorn & Hogarth, 1978; Einhorn, 1980).

(3) In order to understand the relations between predictions, actions, and outcomes, one needs a causal model of the process. Such a model must be developed through the use of what we call diagnostic or backward inference. That is, past observations, events, and data are used as evidence to infer the process(es) that produced them. Such inferences are called diagnostic since they involve going from visible effects such as symptoms, signs, and the like, to their prior causes. For example, one might consider a particularly large decline in last month's sales as symptomatic of a more fundamental malady (such as an incompetent sales force). We consider diagnostic inference to be based on causal thinking, although in doing diagnosis one has to mentally reverse the time order in which events were thought to have occurred (hence the term "backward inference"). On the other hand, predictions involve forward inference; i.e., one goes forward in time from present causes to future effects. However, it is important to recognize the dependence of forward inference/prediction on backward inference/diagnosis. In particular, it seems likely that success in predicting the future depends to a considerable degree on making sense of the past. Therefore, people are continually engaged in shifting between forward and backward inference in both making and evaluating forecasts. Indeed, this can be eloquently summarized by Kierkegaard's observation that, "Life can only be understood backwards; but it must be lived forwards."

A second important aspect of diagnostic inference concerns the process by which hypotheses are formed and relevant variables found. For example, why are certain variables chosen as "relevant" to the phenomenon in question while others are considered irrelevant or of lesser importance? This issue goes to the heart of all prediction problems although it has not received the attention it deserves. For example, consider the clinical vs. statistical prediction controversy in psychology, which has raged off and on for over 25 years (Meehl, 1954; Sawyer, 1966, Dawes, 1979). One succinct conclusion of this literature (and one that we subscribe to) has been made by Dawes and Corrigan (1974); namely, ". . . the whole trick is to decide what variables to look at and then to know how to add" (1974, p. 105). Assuming that we can add, <u>how</u> do we decide what variables to look at? On this crucial point we have little guidance other than some ill-defined notions of experience, intuition, gut feeling, hunch, and so forth. From our perspective, one must have some hypothesis or theory for selecting relevant from irrelevant variables. Indeed, relevance can only be understood in relation to some model (usually implicit) of what generates the variable to be predicted. For example, imagine being asked to predict the result of mixing two chemicals with known atomic structures. Without some knowledge of chemistry and physics, picking relevant from irrelevant variables is meaningless. Therefore, prediction depends on backward inference which involves both the forming of hypotheses to interpret the past and the choosing of relevant from irrelevant variables in that interpretation.

We now turn to consider the details of the diagnostic process with particular emphasis on how people seek relevant variables and test for causal significance. These issues are examined in the following way: (1) We discuss the role of similarity and differences in directing attention to variables; (2) We posit the existence of probabilistic signs used for inferring causal relations and call these "cues to causality;" (3) We identify the following cues and discuss their roles in causal thinking:

temporal order, constant conjunction, contiguity in time and space, number of alternative explanations, similarity, predictive validity, and robustness; (4) The role of the cues in inferring causality is emphasized by discussing the nature of spurious correlation and the existence of causal relations in the absence of correlation; (5) We conclude by considering implications for forecasting.

The Diagnostic Process

We conceive of diagnosis as consisting of three inter-related phases: (1) finding relevant variables; (2) linking variables into causal chains; and (3) testing the causal *significance* of the links in those chains. In this paper we limit ourselves to how two, or at most, three variables become linked in a causal manner. We begin by first discussing the general issues of what directs attention to variables and what is meant by "causal relevance."

Much psychological research indicates that processes of perception and judgment are sensitive to differences or deviations from present states, adaptation levels, and reference points (e.g., Helson, 1964; Kahneman & Tversky, 1979). Therefore, we propose that in *searching* for a cause of some effect, which is itself a deviation from the normal or average, attention is directed toward prior deviations or abnormal states of comparable length and strength. For example, our cave-dwellers probably asked themselves what unusual event (or events) preceded the decline in births. Moreover, if the effect of interest is large (i.e., is of substantial duration and/or strength), we expect that the suspected cause(s) are judged to be of comparable size. These conjectures follow from what Kahneman and Tversky call the "representativeness" heuristic (1972), which in this case implies that similar causes have similar effects. Indeed, Mill noted that this is a deeply rooted belief that, "not only reigned supreme in the ancient world, but still possesses almost undisputed dominion over many of the most cultivated minds" (cited in Nisbett & Ross, 1980, p. 115). Mill thought that such a belief was erroneous and many cases exist in which similarity has been misleading (see e.g., Shapiro, 1960; Shweder, 1977). On the other hand, it is difficult to imagine how one could search for variables without using some notion of similarity. Of particular importance is the non-literal use of similarity through analogy and metaphor (e.g., Ortony, 1979). Indeed, analogies and metaphors provide models of phenomena and thus direct attention to specific aspects and variables. Moreover, their use engages one's prior knowledge so that the explanation for a new or poorly understood phenomenon can be integrated with what is already known or believed. For example, in trying to understand how the brain works, one could consider it as a computer, a muscle, or a sponge. Note how the metaphors direct attention to different, but known, variables and processes. Thus, a computer model suggests informational input, retrieval, and computational processes; a muscle model suggests that processes are strengthened and weakened with amount of use, that thinking can be a "strain" (cf. Shugan, 1980), and so on; a sponge model suggests a more passive "soaking up of information," etc. As should be clear, the choice of a particular metaphor is crucial since it directs attention to a limited set of variables, thereby excluding others. However, this is precisely the function it should serve (we consider this further in the "Implications" section).

While analogies and metaphors help direct attention to variables, not all variables considered are seen as <u>causally</u> relevant. Therefore, it is important to delineate how this occurs. In order to do so, we discuss the concept of a "causal field" and a causal variable as a difference-in-that-field (Mackie, 1974). The following example illustrates both concepts:

Imagine a worker in a chemical plant who contracts cancer and sues the company for causing his disease. His lawyer argues that the cancer rate of workers in this factory is nine times the national average for workers in comparable industries. Note that the "field" in this argument is industries of a certain type and the causal argument rests on a difference (higher cancer rates) in this field. However, the defense lawyer asks: "Would the worker have gotten cancer if he didn't work in the chemical plant?" In order to answer this, he then shows that there is a history of cancer in the worker's family (none of whom worked in chemical factories); the worker has smoked cigarettes for 20 years prior to getting cancer; and so on. The important point to note is that the field has now shifted to people who have certain family histories and habits; and in this field, cancer is not unusual and thus not a difference in the field. Hence, this particular case of cancer is not causally related to working in chemical factories. Of course, one could strengthen the prosecution case if the field were again changed--e.g., showing that the particular chemical plant had nine times the cancer rate of other chemical plants making the same products. The reason that this information strengthens the prosecution's case is that by narrowing the field to chemical plants making the same products, the possibility of alternative explanations is reduced thereby making the difference in the narrowed field more causally relevant.

Although people are sensitive to differences-in-a-field, there are many differences to which one could attend. Therefore, what mechanisms guide the allocation of attention to those variables of greatest causal significance? An important consideration in this regard concerns the factors that distinguish between differences that are causal vs. those that are called "conditions." For example, imagine a house fire in which a gasoline can was found. While the flammable material is a difference in the field (since most people do not have gasolines cans in their homes), it is not likely to be seen as the cause of the fire. Indeed, one is more likely to consider the throwing of a lighted match or cigarette near the can as the cause. As Mackie (1974) points out, there are no general rules for distinguishing causes from conditions, although the following are useful: (i) events are more causal than standing conditions (sparks rather than flammable material cause fires); (ii) events that are intrusive are more causal than those that generally occur; (iii) something abnormal or wrong is more causal than what is normal and right (e.g., the accident was caused by the person veering to the left--not by the other person who drove straight ahead). It is important to stress that there are exceptions to all three rules; e.g., poverty, which is a standing condition, may be seen as the cause of social unrest. Therefore, each of these rules only provides probabilistic evidence for distinguishing causes from conditions.

The mechanism by which causal links are formed and tested is now considered by introducing the concept of "cues to causality." We consider the term "cues" as having a specific meaning that corresponds with its use in Brunswik's psychology (1952; see also Hammond, 1955; Campbell, 1966). Specifically: (1) The relation between each cue and causality is probabilistic. That is, each cue is only a fallible sign of a causal relation; (2) People learn to use multiple cues in making inferences in order to mitigate the potential errors arising from the use of single cues; (3) The use of multiple cues is facilitated by the intercorrelation (redundancy) between cues in the environment. Such intercorrelation reduces the negative effects of omitting cues and aids in directing attention to the presence of other cues; (4) Although multiple cues reduce uncertainty in inference, they do not entirely eliminate it . This is as true for backward as it is for forward inference.

In order to provide a conceptual framework for discussing the cues, imagine that one has a causal candidate in mind (X) for explaining some effect, Y. The causal strength of X will depend on its being a difference-in-the-field, which can be assessed by considering it relative to Y <u>and</u> $\bar{Y}$; i.e., does X discriminate between the occurrence or absence of Y. For example, in trying to determine why some people got sick after eating in a particular restaurant (Y), we immediately want to know if those who didn't get sick ($\bar{Y}$), ate the same food (X). If the number of cases of $X \cap \bar{Y}$ and $X \cap Y$ are comparable, the strength of the connection is clearly diminished. However, this is not the only difference that can be considered. One could further ask whether Y is a difference-in-the-field of X vs. $\bar{X}$; i.e., the field is reversed so that what is figure becomes ground, and vice versa. In this case we wish to compare the number of people who got sick <u>without</u> eating the suspected food ($\bar{X}$), to those who didn't get sick ($\bar{Y} \cap \bar{X}$). These differences, as well as those above, can be summarized in a standard 2 x 2 table, as shown in Table 1. The frequencies in each of the four cells can be thought of as individual

<u>Insert Table 1 about here</u>

components that are combined in various ways to judge the relation between X and Y. The cues to causality come into play by directing attention to different parts of the 2 x 2 table. We now consider the cues.

(1) The cue of <u>temporal order</u> has two related functions; (a) since causes are assumed to precede effects, the search for relevant causal variables focuses on only those events that preceded the effect and thereby reduces the number of variables to be considered; (b) in terms of Table 1, temporal order is used to label the axes of the 2 x 2 table as being either X (independent variable) or Y (dependent variable).

(2) The cue of constant conjunction is defined as the degree to which two variables occur together, holding contiguity in time and space constant. For example, consider a specific case of classical conditioning where a bell is repeatedly paired with food such that an animal learns to salivate at the sound of the bell. However, when the bell is not always followed by food, the relation is more difficult to learn. Note that temporal order is also involved in this example; in fact, attempts at "backward conditioning" (giving food and then the bell) have generally been unsuccessful (however, see Spetch, et al., 1981). In terms of Table 1, we consider that constant conjunction directs attention to the joint frequency of X and Y (cell a).

(3) We treat the cues of contiguity in time and space together although they are conceptually distinct. The reason for considering them together is that they both influence the relevance of constant conjunction as a causal cue and thus concern cell a in Table 1; i.e., contiguity is viewed as a cue to causality via its effects on constant conjunction. This occurs in two ways: (a) Together with similarity, attending to differences-in-a-field, and temporal order, contiguity in time and space aids in focusing attention on what variables occurred close in time to, and in the vicinity of, some effect Y. Indeed, Siegler has shown that for young children (5-6 years old), temporal contiguity is a very strong cue for inferring causality (Siegler & Liebert, 1974; Siegler, 1976). Moreover, these studies show that older children are less dependent on contiguity alone, being able to make use of multiple cues. Nevertheless, in the absence of high contiguity, variables that are causally related may not be noticed as important, as in the low temporal contiguity between pregnancy and intercourse. Additionally, the importance of contiguity cues can be seen by noting that certain causal distinctions are made when contiguity conflicts with other cues to

causality. For example, consider the distinction between a "precipitating" and "underlying" cause. The former is generally some action or event that is high in temporal and spatial contiguity but low in similarity of length or strength with the effect. Thus, the precipitating cause of World War I was an assassination in Sarajevo but the underlying cause(s) were economic upheaval, German nationalism, and so on. (b) Once variables have been identified, we hypothesize that constant conjunction is weighted by the degree of temporal and spatial contiguity between the variables. Therefore, although two variables may have a high joint frequency of occurrence, we would expect their judged causal strength to be lowered as temporal and spatial contiguity were decreased. Similarly, we expect that increased contiguity in time and space will strengthen the causal relevance of variables (including those with low joint frequency).

(4) The <u>number of alternative explanations</u>, or competing variables as causes of Y, is an important negative cue to causality (i.e., the greater the number, the less the causal relevance of some X for Y). As stressed by Campbell and colleagues (Campbell & Stanley, 1963; Campbell, 1975; Cook & Campbell, 1979), causal strength should be evaluated by the ruling out of alternative explanations. Indeed, Mackie (1974) states that the primitive notion of a cause involves asking oneself the question: "Would Y have occurred if X had not?" The greater the number of alternative explanations underlying a "yes" answer, the lower the causal relevance of X for Y. Note that the posing and answering of the above question (which is called a "counterfactual conditional") may either involve doing a real experiment or what is called a "thought" experiment. In the former, one compares the effect of X on Y with that of $\bar{X}$ on Y (the control group condition). In this way, the counterfactual question is easily answered. Moreover, note that a control

group allows one to infer that X is the only difference in the field that can affect Y (since the control group *is* the field). However, even in real or quasi-experiments, alternative explanations can exist if other cues to causality are ambiguous. For example, Campbell (1969) points out that: (a) the gradual introduction of some experimental change makes the determination of its causal impact more difficult than a sharp introduction, as for example when a remedial program is phased in over a long time period rather than implemented all at once; (b) unless replication of the $X \cap Y$ relation is accomplished by the introduction of X over multiple time periods or over different units within the same time period, evidence of constant conjunction is weak; (c) the determination of causal relations when effects are not contiguous in time and space with the manipulated variables is problematic. For example, variables that have "lagged" effects or cumulative effects over time are difficult to isolate.

When real or quasi-experiments are not possible, one can nevertheless engage in the following thought experiment in order to answer the question, "Would Y have occurred if X hadn't?": Imagine the world before X, go forward to where X would occur, and then delete it from the scenario. Now run the scenario forward from that point to see if Y happens or not. Clearly, in such thought experiments the construction of "possible worlds" and imaginary scenarios is crucial for judging causal significance. In fact, according to Mackie (1974),

> The key item is a picture of what *would* have happened if things had been otherwise, and this is borrowed from some experience where things *were* otherwise. It is a contrast case rather than the repetition of like instances that contributes most to our primitive concept of causation. (p. 57)

The use of alternative explanations as a cue to causality has important implications for making inferences in general and for interpreting experiments in particular. As recognized by Hume, "not only in philosophy, but even in common life, we may attain the knowledge of a particular cause merely by one experiment, and after a careful removal of all foreign and superfluous circumstances" (as quoted in Mackie, 1974, p. 25). As a case in point, consider the following one-shot case study with a single datum: The occurrence of a huge explosion near Los Alamos, New Mexico, in July 1945 which no one doubted to be the effect of detonating an atomic bomb. Clearly, inferring causality in this poorly designed experiment was not difficult whereas assessing causality in the most meticulously designed experiments in social science is often problematic at best. When one considers why the causal inference is so strong in the bomb example, ask yourself the following question: "Would an explosion of such magnitude have occurred if an atomic bomb had not gone off?" While it is possible to think of alternative explanations for the explosion, they are so unlikely as to be virtually non-existent. Moreover, note how the other cues to causality point to a causal relation (temporal order, contiguity in time and space, similarity of the "unusualness" of effect and cause, and so on). Therefore, even in one-shot case studies with no control group, causality can be inferred (see Campbell, 1975 for an illuminating discussion of this issue).

The use of counterfactual questions for assessing alternative explanations can be conceptualized by referring back to Table 1. Note that in backward or diagnostic inference, the question, "Would Y have occurred if X hadn't?," focuses attention on cell c, which contains the $Y \cap \bar{X}$ instances. In forward inference, on the other hand, one can ask the question, "Would not-Y occur if X had?". In this case, attention is focused on cell b, containing the $\bar{Y} \cap X$ cases. In either situation, one moves from considering only cell a in Table 1 to other components of the X,Y relation.

(5) As noted earlier, <u>similarity</u> plays a crucial role in the finding of relevant variables. Furthermore, analogy and metaphor can be used to under-

stand new phenomena by linking them to one's prior knowledge. However, one can ask how good a certain analogy might be, as for example, between the human brain and the computer. Whereas both possess many common features, they also contain distinctive aspects. Tversky (1977) has proposed a theory of how such common and distinctive aspects are combined to form similarity judgments. Specifically, he posits that the judgment of similarity between two objects is a weighted linear function of the features they have in common less the distinctive features of each. The parameters of this function reflect differential attention paid to common and distinctive components. Tversky's formulation parallels the judgment of causal strength in the following way. Consider Table 1 and note that the frequencies in cells a and d can be seen as representing the constant conjunction of the variables (analogous to common features), whereas the frequencies in cells b and c can be seen as instances that disconfirm constant conjunction or affirm alternative explanations (analogous to the distinctive feature of two objects). If one considers how attention can be shifted to these different aspects by such factors as whether the task involves forward or backward inference, or what other cues to causality are present (temporal order, contiguity), then judged causal strength can be seen as a weighted linear function of confirming and disconfirming evidence (for experimental results, see Schustack & Sternberg, 1981).

(6) The degree to which one variable can predict another, denoted as <u>predictive validity</u>, is an important cue to causality. Predictive validity is commonly measured by the correlation coefficient and it is noteworthy that this involves all four cells of Table 1. That is, the correlation between X and Y, denoted as r_{xy}, is given by the formula, $r_{xy} = (ad-bc)/[(a+b)(c+d)(a+c)(b+d)]^{1/2}$. Other things being equal, the greater the predictive validity of some variable, the greater its causal relevance. However, when other

causal cues are involved, things are not equal. This makes the use of statistical correlation as a causal cue more uncertain than is generally realized. We discuss this in some detail by considering the meaning of "spurious correlation." Thereafter, we consider what we call "causalation," which results when variables that are causally related show low or no statistical correlation.

On the psychology of spurious correlation. Every student who has taken an introductory statistics course containing a section on correlation is told: correlation does not imply causation. Unfortunately, the factors that do imply causation are rarely if ever discussed and students are left without further guidance. To make matters worse, several days later the same students are warned against "spurious correlation," i.e., the correlation between two variables due to the common causal influence of some third factor. Since the concept of spurious correlation suggests that some correlations are more (or less) causally related than others, it is natural to ask how one can tell the difference (cf. Simon, 1954). However, one is disappointed to learn that there are no simple rules for doing this and that one must exercise one's judgment in this matter. This is not to say that such judgments will always be difficult to make or that they will vary according to the person doing the judging. For example, consider the correlation between the number of pigs and the amount of pig-iron (Ehrenberg, 1975). Such a correlation does seem spurious when the common causal factor, "economic activity," is considered. On the other hand, consider the correlation between amount of rain and number of auto traffic accidents in a city, over the course of a year. Such a correlation does not seem spurious. What is the difference between these two cases?

If we make use of the cues to causality, the spuriousness of the correlation between pigs and pig-iron becomes apparent. First, consider the cues that do point to a causal relation; viz., constant conjunction, contiguity in time, and predictive validity. However, the cues that point away from a causal relation are: temporal order (which cannot be used to specify which variable is cause or effect); low contiguity in space (it being unlikely that farms and factories are in close physical proximity); many alternative explanations for either variable (for example, the answer to the question, "Would pig production have gone up if the production of pig-iron hadn't?," has many yes answers); the similarity of the variables is only with respect to their names, all else being quite dissimilar; the robustness of predictive validity seems low when, for example, one considers the lack of relation in non-industrialized countries. Taken together, the negative evidence regarding a causal relation seems much stronger than the positive evidence. Indeed, the judgment that the relation is spurious is made easily and quickly.

Now consider the second case: the temporal order of rain and accidents is clear; constant conjunction, contiguity in time and space, and predictive validity are all high; similarity, via the use of prior knowledge about the effects of slippery roads, is high; robustness is also high since predictive validity holds in many cities over widely dispersed geographical locations. The only negative cue is the number of alternative explanations since the answer to the question, "Would accidents have increased if it hadn't rained?", might be yes. However, even in this case, there may be few competing alternatives. Therefore, taken together, the cues strongly point to a causal relation and a judgment that the correlation is non-spurious seems warranted.

The generalization that can be made from the above is the following: the judged strength between variables is a joint function of their correlation and the causal cues that are implicit in the labeling of the variables. This statement has several important implications: (1) The labeling of variables should have a major effect on their judged relational strength, holding statistical correlation constant. Evidence for this comes from a study by Jennings, Amabile, and Ross (in press). They found that when people viewed scatterplots of variables labeled X and Y, the statistical correlation had to be quite high for people to see a relationship. However, when the variables were given labels that engaged prior knowledge (making use of the cues to causality), the statistical correlation needed to see a relationship was much lower; (2) The effects of labeling have also been studied in probability learning studies and confirm the importance of cues to causality in learning the relations between variables. For example, Adelman (1981) found that subjects in a multiple-cue probability task learned quite well when variable labels were congruent with statistical predictiveness, but not otherwise. Camerer (1981) showed that subjects were able to learn a disordinal interaction only when the variables were labeled in accord with prior beliefs (this involved factors that were thought to affect the price of wheat futures in a commodity market). When the same task was given as an abstract problem with variables labeled as X_1, X_2, and Y, no learning occurred; (3) When statistical correlation and cues to causality conflict, spurious correlation is not the only outcome; e.g., a low or zero statistical correlation could mask a true causal relation. In order to illustrate this, let us return to our cave-dwellers, who have hit upon the hypothesis that sexual intercourse is related to pregnancy. Accordingly, they designed and carried out the following experiment: One hundred females were allocated at random to an intercourse condition, and 100 to a non-intercourse condition. As indicated in Table 2, 25 females became pregnant, and 175 did not. Given our present

Insert Table 2 about here

world knowledge we can state that the 5 people in the no-intercourse/yes-pregnancy cell represent "measurement error," i.e., faulty memory in reporting, lying, etc. Is intercourse important for pregnancy? The statistical correlation is small ($r = .34$) and our cave-dwellers, in their methodological sophistication, might well question whether the hypothesis is worth pursuing. Indeed, if the sample size were smaller, the correlation might not even be statistically "significant". Moreover, even with a significant correlation, $r^2 = .12$, which is hardly a compelling percentage of the Y variance accounted for by X.

The important implication of the above hypothetical experiment is the following: although correlation doesn't necessarily imply causation, causation doesn't necessarily imply correlation. We have somewhat facetiously labeled examples of the latter as "causalations," giving it equal standing with the better-known and opposite concept of spurious correlation. The importance of causalation is that it demonstrates that sole reliance on statistical measures for understanding and interpreting data is insufficient (see also Simon, 1954). This conclusion is not surprising to those who have always maintained that judgment be a part of the evaluation of evidence. However, the delineation of the cues to causality gives one some hint as to the components of such judgments.

(7) We have used the conceptual device of Table 1 to show how the various cues to causality direct attention to different aspects of causal strength. However, the cue of robustness goes beyond a single 2 x 2 table and explicitly raises the question as to whether, and to what degree, the predic-

tive validity between X and Y varies as a function of other variables (see Toda, 1977). For example, imagine that there is a positive correlation between smoking (X and $\bar{X}$) and lung cancer (Y and $\bar{Y}$). Now consider that the correlation is computed separately for men and women with the following result: the correlation is positive for men but negative for women (i.e., women who don't smoke are more likely than smokers to get lung cancer). Note that by sub-dividing the original sample into sub-groups, one now considers several 2 x 2 tables and asks whether the overall correlation holds in each. If it doesn't, the relation is not robust and the causal relevance of X for Y is decreased. In the above example, one is likely to be suspicious of any causal relation if the sign of the correlation changes in one of the sub-groups. On the other hand, if predictive validity is robust, it points more strongly to a causal relation.

Implications

We now consider three implications of the above analysis for forecasting: (1) the selection of variables and the building of forecasting models; (2) the evaluation of forecast accuracy; and (3) learning to improve forecasting ability.

(1) We have emphasized the role of similarities and differences as well as the cues to causality in directing attention to variables in the diagnostic process. However, a significant feature of causal/diagnostic thinking is the remarkable speed and fluency which people seem to have for generating explanations and accommodating discrepant facts into expanded hypotheses (Fischhoff, 1975; Slovic & Fischoff, 1977; Tversky & Kahneman, 1980). A useful analogy of this process is provided by multiple regression analysis. In deriving a model (backward inference) one seeks the combination of variables and parameters

that maximizes some measure of fit, e.g., R^2. Furthermore, by increasing variables, R^2 approaches 1. However, and often to the user's surprise, the measure of fit typically "shrinks" on prediction. Thus, the power of post-hoc explanations is unfortunately matched by the paucity of predictive validity.

Since diagnostic thinking is so fluent, one must guard against the way cues to causality quickly restrict our interpretation of the past. Of particular significance in this regard is the use of metaphors in guiding attention and providing models of complex phenomena. Since metaphors, like other cues to causality, are of imperfect validity, errors can be made by relying on single images. "The map," as the general semanticists remind us, "is not the territory" and thus we specifically recommend the <u>method of multiple metaphors</u> as a way of guarding against the premature adoption of a single model. That is, instead of focusing on a single metaphor, experiment with several. For example, consider how one might view forecasting. In this paper we have adopted a medical metaphor by focusing on the diagnostic process and the effects of treatments (actions) on outcomes. However, consider a ballistics metaphor in which forecasting is likened to aiming and shooting at a target (see Hogarth, 1981). Such a metaphor leads to consideration of quite different variables and issues. Therefore, while no single model is correct, each directs attention to different factors, thereby providing a more complete picture of the phenomenon.

Whereas the cues to causality play an important role in structuring and stabilizing our perceptions of reality, such stability may be purchased at the cost of novelty and originality. That is, the cues direct attention to what is "obvious," thereby reducing innovation and creativity. This suggests that one way to facilitate creative thinking might be to go specifically against the cues. For example, when dealing with a complex outcome, one might search for a dissimilar and simple causal candidate rather than a similar and complex one.

(2) As emphasized at the beginning of this paper, the evaluation of forecast accuracy is problematic without a causal understanding of the factors that influence outcomes. In particular, we stressed the fact that forecasts are made to aid decision making and this makes outcomes a function of actions as well as predictions. The difficulties of evaluating forecast accuracy under these conditions can be conceptualized by considering two factors: (a) To what extent are people aware that their actions can affect outcomes? and; (b) to what extent do their actions actually affect outcomes? If, for the sake of simplicity, we only consider two levels of each factor (i.e., aware-not aware and, actions affect outcomes-don't affect outcomes), Table 3 presents a convenient way to summarize the various possibilities.

Insert Table 3 about here

Cells, 1, 2, and 4 were discussed earlier in our examples of economic forecasting, rumors of bank failure, and hurricane prediction, respectively. We simply note that in the first two cases (cells 1 and 2), the evaluation of forecast accuracy should be conditioned on the action taken. That is, let Y be some outcome, X a forecast, and A an action based on the forecast; one is then interested in $p(Y|X,A)$ rather than $p(Y|X)$. As we have noted elsewhere (Einhorn & Hogarth, 1978), when actions affect outcomes, forecasts and outcomes are not conditionally independent of action; i.e., $p(Y|X,A) \neq p(Y|X)$.

The situation represented in cell 3 deserves special attention since it involves taking action to affect outcomes when such actions are useless. When action involves physical manipulation or direct intervention, Langer (1975) has shown that people are prey to an "illusion of control;" i.e., actions are seen to affect outcomes that are generated by random processes (cf. Lopes, in press). For example, people tend to believe that the lottery tickets they personally select have a greater chance of winning than those selected for them by lottery administrators. Similarly, the cognitive activities of planning and forecasting in organizations can lead to illusions of control and overconfidence by restricing attention to those consequences one can imagine while diverting attention from those not considered (Hogarth & Makridakis, 1981). Moreover, it should be noted that illusions of control are conceptually identical to certain types of superstitious behavior. Indeed, the joint presence of randomness and the cues to causality (especially temporal order, constant conjunction, and high contiguity in time and space) inevitably lead to some superstitious behavior (Skinner, 1966). Thus, one can consider superstition as the cost one pays to gain causal knowledge. The issue remains, of course, as to whether the cost outweighs the associated benefits (cf. Killeen, 1978).

(3) Given the difficulties of evaluating forecast accuracy, what can be done to improve forecasting ability? We consider two related approaches: (a) experimentation to untangle the forecast-action-outcome sequence; and (b) systematizing thought experiments.

(a) Opportunities for experimentation are often overlooked. For example, consider personnel selection or advertising. In both cases one could carry out full or partial experiments by randomly selecting employees or stopping advertising completely. While such experiments are typically infeasible, partial experiments could provide much useful information. For example, one could randomly admit a small percentage of candidates and advertising could be stopped in selected time periods or areas (see Cook & Campbell, 1979, for relevant experimental designs).

(b) To systematize thought experiments, we suggest the following procedure. Imagine that sales have increased and you wish to determine the causal significance of a prior advertising campaign. A real experiment would involve using a control group to test whether no campaign would have led to the same level of sales. Using the symbols of Table 1, the control group outcome would focus on the $\bar{X}\cap Y$ cell (i.e., c). Moreover, in the *absence* of a real experiment this datum would be supplied by the answer to the counterfactual question "Would Y have occurred if X hadn't?" However, answering only this question would be a weak form of inference. We also suggest that people ask themselves questions that illuminate the $X\cap\bar{Y}$ and $\bar{X}\cap\bar{Y}$ cells. The former involves imagining whether no sales increase ($\bar{Y}$) would have followed the advertising campaign (X); the latter asks whether no campaign ($\bar{X}$) would have been followed by no increase in sales ($\bar{Y}$). Systematically posing these different counterfactual questions can increase the power of thought experiments by generating information analogous to that available from real experiments.

Conclusion

Forecasting depends on the use of data and the exercise of judgment. Indeed, a colleague has succinctly summarized the theme of this paper by saying that one cannot understand statistics without psychology (see also Hogarth, 1975). While agreeing, we also note that the meaning and effects of uncertainty are central to the psychology of thinking and thus call for statistical expertise. In this paper we have emphasized the probabilistic nature of cues to causality and the uncertainties associated with backward and forward inference. Moreover, we have stressed the need to understand how judgment affects the generation and testing of formal statistical models. We believe that the development and use of statistical techniques will benefit from understanding judgmental processes. Without such understanding, which depends crucially on causal thinking, simply forecasting the future will not ensure the future of forecasting.

References

Adelman, Leonard, (1981), The influence of formal, substantive, and contextual task properties on the relative effectiveness of different forms of feedback in multiple-cue probability learning tasks. Organizational Behavior and Human Performance, 27, 423-442.

Brunswik, Egon, (1952), Conceptual framework of psychology. Chicago: University of Chicago Press.

Camerer, Colin F., (1981), The validity and utility of expert judgment. Unpublished Ph.D. dissertation, University of Chicago.

Campbell, Donald T., (1966), Pattern matching as an essential in distal knowing. In K. R. Hammond (Ed.), The psychology of Egon Brunswik. New York: Holt, Rinehart and Winston.

Campbell, Donald T., (1969), Reforms as experiments. American Psychologist, 24, 409-429.

Campbell, Donald T., (1975), "Degrees of freedom" and the case study. Comparative Political Studies, 8, 178-193.

Campbell, Donald T. and Stanley, Julian C., (1963), Experimental and quasi-experimental designs for research. Chicago: Rand-McNally.

Cook, Thomas D. and Campbell, Donald T., (1979), Quasi-experimentation: Design & analysis for field settings. Chicago: Rand-McNally.

Dawes, Robyn M., (1979), The robust beauty of improper linear models in decision making. American Psychologist, 34, 571-582.

Dawes, Robyn M. and Corrigan, Bernard, (1974), Linear models in decision making. Psychological Bulletin, 81, 95-106.

Ehrenberg, Andrew, S. C., (1975), Data reduction: Analyzing and interpreting statistical data. New York: Wiley.

Einhorn, Hillel J., (1980), Learning from experience and suboptimal rules in decision making. In T. S. Wallsten (Ed.), Cognitive processes in choice and decision behavior. Hillsdale: Erlbaum.

Einhorn, Hillel J. and Hogarth, Robin M., (1978), Confidence in judgment: Persistence of the illusion of validity. Psychological Review, 85, 395-416.

Fischhoff, Baruch, (1975), Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1, 288-299.

Hammond, Kenneth R., (1955), Probabilistic functionalism and the clinical method. Psychological Review, 62, 255-262.

Helson, Harry, (1964), Adaptation-level theory. New York: Harper.

Hogarth, Robin M., (1975), Cognitive processes and the assessment of subjective probability distributions. Journal of the American Statistical Association, 70, 271-289.

Hogarth, Robin M., (1981), Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. Psychological Bulletin, 90, 197-217.

Hogarth, Robin M. and Makridakis, Spyros, (1981), Forecasting and planning: An evaluation. Management Science, 27, 115-138.

Howard, Ronald A., Matheson, James E. and North, D. Warner, (1972), The decision to seed hurricanes. Science, 176, 1191-1202.

Jennings, D., Amabile, T. M. and Ross, L., (in press), Informal covariation assessment: Data-based vs. theory-based judgments. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press.

Kahneman, Daniel and Tversky, Amos, (1972), Subjective probability: A judgment of representativeness. Cognitive Psychology, 3, 430-454.

Kahneman, Daniel and Tversky, Amos, (1979), Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.

Killeen, Peter T., (1975), Superstition: A matter of bias, not detectability. Science, 199, 88-90.

Langer, Ellen J., (1975), The illusion of control. Journal of Personality and Social Psychology, 32, 311-328.

Lopes, Lola L., (in press), Doing the impossible: A note on induction and the experience of randomness. Journal of Experimental Psychology: Human Learning and Memory.

Mackie, John L., (1974), The cement of the universe: A study of causation. Oxford: Clarendon Press.

Meehl, Paul E., (1954), Clinical versus statistical prediction: A theoretical analysis and review of the literature. Minneapolis: University of Minnesota Press.

Nisbett, Richard E., and Ross, Lee D., (1980), Human inference: Strategies and shortcomings of social judgment. Englewood Cliffs: Prentice-Hall.

Ortony, Andrew, (1979), Beyond literal similarity. Psychological Review, 86, 161-180.

Sawyer, Jack, (1966), Measurement and prediction, clinical and statistical. Psychological Bulletin, 66, 178-200.

Schustack, Miriam W. and Sternberg, Robert J., (1981), Evaluation of evidence in causal inference. Journal of Experimental Psychology: General, 110, 101-120.

Shapiro, Arthur K., (1960), A contribution to the history of the placebo effect. Behavioral Science, 5, 109-135.

Shugan, Steven M., (1980), The cost of thinking. Journal of Consumer Research, 7, 99-11.

Shweder, Richard A., (1977), Likeness and likelihood in everyday thought: Magical thinking in judgments about personality. Current Anthropology, 18, 637-658.

Siegler, Robert S., (1976), The effects of simple necessity and sufficiency relationships on children's causal inferences. Child Development, 47, 1058-1063.

Siegler, Robert S. and Liebert, Robert M., (1974), Effects of contiguity, regularity, and age on children's causal inferences. Developmental Psychology, 10, 574-579.

Simon, Herbert A., (1954), Spurious correlation: A causal interpretation. Journal of the American Statistical Association, 49, 467-479.

Skinner, Burrhus F., (1966), The phylogeny and ontogeny of behavior. Science, 153, 1205-1213.

Slovic, Paul and Fischhoff, Baruch, (1977), On the psychology of experimental surprises. Journal of Experimental Psychology: Human Perception and Performance, 3, 544-551.

Spetch, Marcia L., Wilkie, Donald M. and Pinel, J. P. J., (1981), Backward conditioning: A reevaluation of the empirical evidence. Psychological Bulletin, 89, 163-175.

Toda, Masanao, (1977), Causality, conditional probability and control. In A. Aykac and C. Brumat (Eds.), New developments in the applications of Bayesian methods. Amsterdam: North Holland.

Tversky, Amos, (1977), Features of similarity. Psychological Review, 84, 327-352.

Tversky, Amos and Kahneman, Daniel, (1980), Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), Progress in social psychology (Vol. 1). Hillsdale: Erlbaum.

Table 1

Contingency Table Between X and Y

	Y	$\bar{Y}$
X	Cell a = $f(X \cap Y)$	Cell b = $f(X \cap \bar{Y})$
$\bar{X}$	Cell c = $f(\bar{X} \cap Y)$	Cell d = $f(\bar{X} \cap \bar{Y})$

Note: f stands for frequency.

Table 2

Data Matrix for Hypothetical
Intercourse-Pregnancy Experiment

		Pregnancy: Yes	Pregnancy: No	
Intercourse	Yes	20	80	100
	No	5	95	100
		25	175	200

Table 3

Categorization of Forecasting Situations by Awareness and Efficacy of Actions

Awareness of actions affecting outcomes	Actions affect outcomes: Yes	Actions affect outcomes: No
Yes	(1) Forecast accuracy conditioned on action; $p(Y\|X,A) \neq p(Y\|X)$	(3) Illusions of control; superstitions.
No	(2) Self-fulfilling and self-defeating prophecies	(4) Simple forecast accuracy is adequate.

DISTRIBUTION LIST

OSD

CDR Paul R. Chatelier
Office of the Deputy Under
Secretary of Defense
OUSDRE (E&LS)
Pentagon Room 3D129
Washington, DC 20301

Dr. Stuart Starr
Office of the Assistant Secretary
of Defense (C3I)
Pentagon
Washington, DC 20301

Department of the Navy

Director
Engineering Psychology Programs
Code 455
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217 (5 cys)

Director
Manpower, Personnel and Training
Code 270
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Operations Research Programs
Code 434
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Statistics and Probability Program
Code 436
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Information Systems Program
Code 437
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Code 430B
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Special Assistant for Marine Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Commanding Officer
ONR Eastern/Central Regional Office
ATTN: Dr. J. Lester
Building 114, Section D
666 Summer Street
Boston, MA 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. C. Davis
536 South Clark Street
Chicago, IL 60605

Commanding Officer
ONR Western Regional Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco, CA 96503

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, DC 20375 (6 cys)

Dr. Michael Melich
Communications Sciences Division
Code 7500
Naval Research Laboratory
Washington, DC 20375

Dr. Robert G. Smith
Office of the Chief of Naval Operations
OP987H, Personnel Logistics Plans
Washington, DC 20350

Naval Training Equipment Center
ATTN: Technical Library
Orlando, FL 32813

Human Factors Department
Code N215
Naval Training Equipment Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation Group
Naval Training Equipment Center
Code N-00T
Orlando, FL 32813

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Mr. Warren Lewis
Human Engineering Branch
Code 8231
Naval Ocean Systems Center
San Diego, CA 92152

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, DC 20380

HQS, U.S. Marine Corps
ATTN: CCA40 (MAJOR Pennell)
Washington, DC 20380

Commanding Officer
MCTSSA
Marine Corps Base
Camp Pendleton, CA 92055

Mr. Arnold Rubinstein
Naval Material Command
NAVMAT 0722 - Rm. 508
800 North Quincy Street
Arlington, VA 22217

Mr. Phillip Andrews
Naval Sea Systems Command
NAVSEA 0341
Washington, DC 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 4701
Washington, DC 20360

CDR Thomas Berghage
Naval Health Research Center
San Diego, CA 92152

Dr. George Moeller
Human Factors Engineering Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Dr. James McGrath, Code 302
Navy Personnel Research and Development Center
San Diego, CA 92152

Navy Personnel Research and Development Center
Planning and Appraisal, Code 04
San Diego, CA 92152

Navy Personnel Research and Development Center
Management Systems, Code 303
San Diego, CA 92152

Navy Personnel Research and Development Center
Performance Measurement and Enhancement
Code 309
San Diego, CA 92152

Dr. Julie Hopson
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Mr. Jeffrey Grossman
Human Factors Branch, Code 3152
Naval Weapons Center
China Lake, CA 93555

Human Factors Engineering Branch
Code 1226
Pacific Missile Test Center
Point Mugu, CA 93042

Dean of the Academic Departments
U.S. Naval Academy
Annapolis, MD 21402

CDR W. Moroney
Code 55MP
Naval Postgraduate School
Monterey, CA 93940

Mr. Merlin Malehorn
Office of the Chief of Naval Operations
(OP-115)
Washington, DC 20310

Department of the Army

Mr. J. Barber
HQS, Department of the Army
DAPE-MBR
Washington, DC 20310

Dr. Joseph Zeidner
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Organizations and
Systems Research Laboratory
U.S. Army Research Institute
Alexandria, VA 22333

Technical Director
U.S. Army Human Engineering Labs
Aberdeen Proving Ground, MD 21005

U.S. Army Medical R&D Command
ATTN: CPT Gerald P. Krueger
Ft. Detrick, MD 21701

ARI Field Unit-USAREUR
ATTN: Library
C/O ODCSPER
HQ USAREUR & 7th Army
APO New York 09403

Department of the Air Force

U.S. Air Force Office of Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Air University Library
Maxwell Air Force Base, AL 36112

Dr. Earl Alluisi
Chief Scientist
AFHRL/CCN
Brooks, AFB, TX 78235

Foreign Addressees

North East London Polytechnic
The Charles Myers Road
Livingstone Road, Stratford
London E15 2LJ, ENGLAND

Professor Dr. Carl Graf Hoyos
Institute of Psychology
Technical University
8000 Munich
Arcisstr 21
FEDERAL REPUBLIC OF GERMANY

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Technology Establishment
Teddington, Middlesex TW11 OLN
ENGLAND

Director, Human Factors Wing
Defence & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
CANADA

Dr. A. D. Baddeley, Director
Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF ENGLAND

Other Government Agencies

Defense Technical Information Center
Cameron Station, Building 5
Alexandria, VA 22314 (12 cys)

Dr. Craig Fields, Director
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Dr. Judith Daly
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Professor Douglas E. Hunter
Defense Intelligence School
Washington, DC 20374

Other Organizations

Dr. Robert R. Mackie
Human Factors Research, Inc.
5775 Dawson Avenue
Goleta, CA 93017

Dr. Gary McClelland
Institute of Behavioral Sciences
University of Colorado
Boulder, CO 80309

Dr. Miley Merkhofer
Stanford Research Institute
Decision Analysis Group
Menlo Park, CA 94025

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Professor Judea Pearl
Engineering Systems Department
University of California-Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024

Professor Howard Raiffa
Graduate School of Business Administration
Harvard University
Soldiers Field Road
Boston, MA 02163

Dr. Arthur I. Siegel
Applied Psychological Services, Inc.
404 East Lancaster Street
Wayne, PA 19087

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Amos Tversky
Department of Psychology
Stanford University
Stanford, CA 94305

Dr. Robert T. Hennessy
NAS - National Research Council
JH #819
2101 Constitution Avenue, NW
Washington, DC 20418

Dr. M. G. Samet
Perceptonics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, NW
Washington, DC 20036

Dr. Ward Edwards, Director
Social Science Research Institute
University of Southern California
Los Angeles, CA 90007

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, OK 73069

Dr. Kenneth Hammond
Institute of Behavioral Science
University of Colorado, Room 201
Boulder, CO 80309

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, NW
Washington, DC 20036 (3 cys)

Dr. Richard W. Pew
Information Sciences Division
Bolt Beranek & Newman Inc.
50 Moulton Street
Cambridge, MA 02138

Mr. Tim Gilbert
The MITRE Corporation
1820 Dolly Madison Boulevard
McLean, VA 22102

Dr. Douglas Towne
University of Southern California
Behavioral Technology Laboratory
3716 South Hope Street
Los Angeles, CA 90007

Dr. John Payne
Graduate School of Business Administration
Duke University
Durham, NC 27706

Dr. Baruch Fischhoff
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Andrew P. Sage
School of Engineering
and Applied Science
University of Virginia
Charlottesville, VA 22901

Dr. Leonard Adelman
Decisions and Designs, Inc.
8400 Westpark Dr., Suite 600
Post Office Box 907
McLean, VA 22101

Dr. Lola Lopes
Department of Psychology
University of Wisconsin
Madison, WI 53706

Mr. Joseph G. Wohl
Alphatech, Inc.
3 New England Industrial Park
Burlington, MA 01803

Dr. Rex Brown
Decision Science Consortium
7700 Leesburg Pike, Suite 721
Falls Church, VA 22043

Dr. Wayne Zachary
Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090